DOCUMENT RESUME

ED 437 406 TM 030 558

AUTHOR	Bailey, Phillip; Onwuegbuzie, Anthony J.; Daley, Christine E.
TITLE	The Validation of Three Scales Measuring Anxiety at Different Stages of the Foreign Language Learning Process: The Input Anxiety Scale, the Processing Anxiety Scale, and the Output Anxiety Scale.
PUB DATE	1999-11-17
NOTE	44p.; Paper presented at the Annual Meeting of the Mid-South Educational Research Association (MSERA), (Point Clear, AL, November 17, 1999).
PUB TYPE	Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	*Anxiety; *College Students; Higher Education; *Measurement Techniques; *Second Language Learning; *Student Attitudes; Validity
IDENTIFIERS	Confirmatory Factor Analysis

ABSTRACT

This study examined the psychometric properties of the Input Anxiety Scale, the Processing Anxiety Scale, and the Output Anxiety Scale, (all developed by P. MacIntyre and R. Gardner, 1994) which measure anxiety at the input, processing, and output stages of the foreign language learning process. These scales were administered to 250 university students. Evidence of structural validity was provided via three separate exploratory factor analyses, which revealed one factor for each scale, explaining between 43% and 45% of the variance in scores. Confirmatory factor analyses revealed that the three scales did not represent either a single unidimensional construct underlying foreign language anxiety or MacIntyre and Gardner's three-stage model of anxiety. However, when some items were removed, the scales confirmed the three-stage model, suggesting modifications to the scales are needed. (Contains 8 tables and 60 references.) (Author/SLD)

Reproductions supplied by EDRS are the best that can be made
from the original document.

Running head: FOREIGN LANGUAGE ANXIETY

ED 437 406

SCOPE OF INTEREST NOTICE
The ERIC Facility has assigned this document for processing to: TM
In our judgment, this document is also of interest to the Clearinghouses noted to the right. Indexing should reflect their special points of view. FL

The Validation of Three Scales Measuring Anxiety at Different Stages of the Foreign Language Learning Process: The Input Anxiety Scale, the Processing Anxiety Scale, and the Output Anxiety Scale

Phillip Bailey

University of Central Arkansas

Anthony J. Onwuegbuzie[1]

Valdosta State University

Christine E. Daley

Muscogee County School District

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
✓ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY
A. Onwuegbuzie
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

[1] Correspondence should be addressed to Anthony Onwuegbuzie, Department of Educational Leadership, College of Education, Valdosta State University, Valdosta, Georgia, 31698, or (TONWUEGB@VALDOSTA.EDU)

Paper presented at the annual meeting of the Mid-South Educational Research Association (MSERA), Point Clear, Alabama, November 17, 1999.

TM030558

BEST COPY AVAILABLE

Abstract

This study examined the psychometric properties of the Input Anxiety Scale, the Processing Anxiety Scale, and the Output Anxiety Scale, which measure anxiety at the input, processing, and output stages of the foreign language learning process. These scales were administered to 258 university students. Evidence of structural validity was provided via three separate exploratory factor analyses, which revealed one factor for each scale, explaining between 43% and 45% of the variance in scores. Confirmatory factor analyses revealed that the three scales did not represent either a single unidimensional construct underlying foreign language anxiety or MacIntyre and Gardner's (1984) three-stage model of anxiety. However, when some items were removed, the scales confirmed the three-stage model, suggesting that modifications to the scales are needed.

The Validation of Three Scales Measuring Anxiety at Different Stages of the Foreign Language Learning Process: The Input Anxiety Scale, the Processing Anxiety Scale, and the Output Anxiety Scale [2]

In the past two decades, foreign language researchers and educators have increasingly focused their attention on foreign language anxiety as among the most important affective predictors of foreign language achievement. Foreign language anxiety is best described as a form of situation-specific anxiety (Horwitz, Horwitz, & Cope, 1986; MacIntyre, 1999). That is, it is neither a trait anxiety, which generally refers to a person's tendency to be anxious, nor is it state anxiety, although it often manifests itself in the physiological signs of the latter, including: perspiration, sweaty palms, dry mouth, muscle contractions and tension, and increases in heart and perspiration rates (Chastain, 1975; Gardner, 1985; Steinberg & Horwitz, 1986). Research has indicated that anxiety is common among foreign language students (Aida, 1994),[3] and that it is associated negatively with language performance (Gardner & MacIntyre, 1993; Madsen, Brown, & Jones, 1991; MacIntyre & Gardner, 1991a, 1991b, 1991c, 1994a),[4] and with student self-ratings of second language proficiency (MacIntyre, Noels, & Clément, 1997). Ganschow and Sparks (1996) suggest that a student's anxiety level in foreign language class may be "an early indicator of basic language problems" (p. 199). In fact, anxiety appears to be one of the best predictors of second language achievement (Ehrman & Oxford, 1995; Gardner, 1985; Horwitz, 1986; MacIntyre & Gardner, 1994a, 1994b; MacIntyre et al., 1997; Onwuegbuzie, Bailey, & Daley, 1999a, 1999b). As such, research into the nature of foreign language anxiety holds great promise for improving language learning in the classroom.

Much research exists examining the correlates of foreign language anxiety.

Most recently, Onwuegbuzie, Bailey, and Daley (in press) found that students with the highest levels of foreign language anxiety tended to have at least one of these characteristics: older, high academic achievers, had never visited a foreign country, had not taken any high school foreign language courses, had low expectations of their overall average for their current language course, had a negative perception of their scholastic competence, and had a negative perception of their self-worth. Also, Bailey, Onwuegbuzie, and Daley (in press) found that students with the highest levels of foreign language anxiety tended to report that (1) they spend too much time on some subjects and not enough time on others; (2) they frequently do not get enough sleep and feel sluggish in class or when studying; (3) they do not try to space their study periods so that they do not become too tired while studying; and (4) they have trouble settling down to work and do not begin studying as soon as they sit down.

Until recently, most researchers have treated foreign language anxiety as a unidimensional construct. However, applying Tobias' model of the effects of anxiety on learning, MacIntyre and Gardner (1994b) have theorized that foreign language anxiety occurs at each of the following three stages of the second language acquisition process: input, processing, and output. Although MacIntyre and Gardner are careful to note that "the term *stages* in Tobias' (1986) model should not be taken to mean that learning occurs in discrete sections" (p. 287), they nonetheless contend that the interdependence of the three stages does not preclude that foreign language anxiety can be conceptualized as occurring at these stages.[5]

According to MacIntyre and Gardner (1994a), anxiety at the input stage (i.e., *input anxiety*) represents the fear experienced by foreign language students when they are initially presented with a new word, phrase, or sentence

in the foreign language. The level of anxiety at this stage is a function of the student's ability to receive, to concentrate on, and to encode external stimuli. Anxiety produced at this stage may reduce the efficacy of input. This may occur when the anxious student's ability to attend to material presented by the instructor diminishes, and nominal stimuli become ineffective due to an inability to represent input internally (Tobias, 1977). Students with high levels of input anxiety typically attend more to task-irrelevant information and material, reducing the capacity to receive input (Onwuegbuzie & Daley, 1996). According to MacIntyre and Gardner (1994a), students with high levels of anxiety at the input stage may ask for their foreign language instructors to repeat sentences more often than do their low-anxious counterparts, or may have to reread material in the foreign language on several occasions in order to compensate for missing or inadequate input.

Anxiety at the processing stage denotes the apprehension experienced when cognitive operations are performed on the external stimuli--that is, when students typically are attempting to organize and to store input. The amount of anxiety involved at this stage appears to depend on the difficulty of the material presented, the extent to which memory is relied upon, and the level of organization of the presented material (Tobias, 1986). According to Tobias (1977), anxiety at this stage can debilitate learning by interfering with the processes that transform the input information and generate a solution to the problem. That is, anxiety may reduce the efficiency with which memory processes are utilized to solve the task. In particular, high levels of processing anxiety may reduce a student's ability to understand messages or to learn new vocabulary items in the foreign language (MacIntyre & Gardner, 1994a).

Finally, anxiety at the output stage encompasses the worry experienced when

students are required to demonstrate their ability to produce previously learned material. In particular, anxiety at this stage involves interference that appears after processing has been completed, but before it has been reproduced effectively as output (Tobias, 1977). Tobias (1977) postulated that output anxiety interferes with the retrieval of previous learning. According to MacIntyre and Gardner (1994a), high levels of anxiety at this stage might hinder students' ability to speak or to write in the foreign language.

MacIntyre and Gardner (1994a) developed three scales to measure anxiety at the input, processing, and output stages. Using students enrolled in foreign language courses at a Canadian university, these researchers found anxiety to be related to overall foreign language achievement at each of the three stages. Although MacIntyre and Gardner (1994a) provide estimates of reliability (i.e., coefficient alpha), and evidence that the three scales are significantly correlated with several other foreign language anxiety scales and a variety of tasks at the three stages in question, to date, no other published study has examined the psychometric properties of these instruments. This was the major purpose of the present study. Also examined was the extent to which these scales adequately measure and reflect the three-stage conceptualization.

Method

Subjects

Participants were 258 college students (67.6% female) from a number of disciplines, who were enrolled in Spanish (n = 157), French (n = 75), German (n = 20), and Japanese (n = 6) introductory, intermediate, and advanced courses at a large university in the mid-southern United States. The subjects were volunteers who received extra course credit and were required to give their consent by signing an informed consent document. Participants represented 43

degree programs from the Colleges of Business Administration, Education, Fine Arts and Communication, Health and Applied Sciences, Liberal Arts, and Natural Sciences and Mathematics. With respect to year of study, participants consisted of first-year students (15.2%), sophomores (19.9%), juniors (30.9%), seniors (31.3%), and graduates (1.6%). Mean age for the sample was 22.8 (*SD* = 6.8). Also, mean grade point average was 3.02 (*SD* = 0.62).

Instruments and Procedure

Participants were administered the Input Anxiety Scale, the Processing Anxiety Scale, and the Output Anxiety Scale. These scales, which were developed by MacIntyre and Gardner (1994a), each contain six 5-point Likert-format items (i.e., 1 = *strongly agree*, 2 = *agree*, 3 = *neutral*, 4 = *disagree*, 5 = *strongly disagree*) that assess how anxious students feel at the input, processing, and output stages, respectively. All negative items were key-reversed before scoring, such that high scores on any of these scales represent high levels of anxiety at the corresponding stage. Sample items for the Input Anxiety Scale include, "I get flustered unless French/Spanish/German/Japanese is spoken very slowly and deliberately" and "I get upset when I read in French/Spanish/German/Japanese because I must read things again and again." Sample items for the Processing Anxiety Scale include, "I am anxious with French/Spanish/German/Japanese because, no matter how hard I try, I have trouble understanding it" and "I feel anxious if French/Spanish/German/Japanese class seems disorganized." Finally, sample items for the Output Anxiety Scale include, "I may know the proper French/Spanish/German/Japanese expression but when I am nervous it just won't come out" and "When I become anxious during a French/Spanish/German/Japanese test, I cannot remember anything I studied." MacIntyre and Gardner (1994a) reported coefficient alpha reliabilities of .78,

.72, and .78, for the Input Scale, the Processing Scale, and the Output Scale, respectively. Additionally, the authors provided evidence of construct validity for these scales via statistically significant correlations between each scale and (1) the French Class Anxiety Scale (Gardner, 1985), which assesses the extent to which respondents feel anxious during French classes; (2) the French Use Anxiety Scale (MacIntyre & Gardner, 1988), which measures the degree to which students feel anxious using French outside the classroom; and (3) the Foreign Language Classroom Anxiety Scale (Horwitz et al., 1986), a global measure of foreign language anxiety. Specifically, these authors reported that the IAS was correlated significantly ($p < .001$) with the French Class Anxiety Scale ($r = .67$), the French Use Anxiety Scale ($r = .64$), and the Foreign Language Classroom Anxiety Scale ($r = .62$); the PAS was correlated significantly ($p < .001$) with the French Class Anxiety Scale ($r = .70$), the French Use Anxiety Scale ($r = .64$), and the Foreign Language Classroom Anxiety Scale ($r = .69$); and the OAS was correlated significantly ($p < .001$) with the French Class Anxiety Scale ($r = .82$), the French Use Anxiety Scale ($r = .72$), and the Foreign Language Classroom Anxiety Scale ($r = .81$).

Results

Reliability

Reliability is the extent to which scores that are generated from an instrument demonstrate consistency (Campbell & Stanley, 1990; Gay, 1999; Kerlinger, 1999). Cronbach's Coefficient Alpha provides information about the degree to which the items in a scale measure similar characteristics (Campbell & Stanley, 1990; Gay, 1999; Kerlinger, 1999). Coefficient Alpha, a measure of internal consistency, was determined for each scale, yielding the following reliability estimates: .72 for the IAS, .73 for the PAS, and .75 for the OAS.

Alpha coefficients reported by MacIntyre and Gardner (1994a) were similar (i.e., .78, .72, and .78, respectively). These two sets of reliability estimates are adequate for affective measures (Nunnally, 1994).

Point Multi-Serial Correlation Alpha Coefficients (PMSCACs) were determined for each item within each of the three scales by deleting one item at a time, and then computing the resulting alpha coefficient.[6] This index helps to assess the extent to which each item contributes to a scale. Any item that has a PMSCAC that is much larger than the overall coefficient alpha for the scale to which it belongs should be excluded, since a relatively large PMSCAC indicates that the corresponding item does not contribute sufficiently to the overall coefficient alpha. The PMSCACs are presented in Tables 1-3. It can be seen from these tables that the PMSCACs ranged from .62 to .74 for the IAS, from .65 to .75 for the PAS, and from .69 to .74 for the OAS. Because these ranges were not substantial, no item appeared to require removal.

Construct-Related Validity

Validity is the extent to which an instrument measures what it is supposed to measure (Campbell & Stanley, 1990; Gay, 1999; Kerlinger, 1999; Nunnally, 1994). Furthermore, construct-related validity is the extent to which an instrument can be interpreted as a meaningful measure of some characteristic or quality (Campbell & Stanley, 1990; Gay, 1999; Kerlinger, 1999). Establishing structural validity is an important step in providing evidence of construct validity. Exploratory factor analysis was used to assess the structural validity of the scales. Specifically, a maximum likelihood (ML) factor analysis was used to determine the number of factors underlying each scale. This technique, which is more valid for identifying the number and nature of the latent factors that are responsible for covariation in a dataset than is principal components factor

analysis' (Bickel & Doksum, 1977; Hatcher, 1994), is perhaps the most commonly used method of common factor analysis (Lawley & Maxwell, 1971). The ML factor analyses, with no constraints imposed, revealed (1) one specific factor for the Input Anxiety Scale, which explained 43.3% of the total variance; (2) one specific factor for the Processing Anxiety Scale, which explained 44.0% of the total variance; and (3) one specific factor for the Output Anxiety Scale, which explained 44.7% of the total variance. Loadings of items on each factor and percent of variance explained are presented in Tables 1-3. It can be seen from these tables that the loadings ranged from .30 to .78 for the IAS, from .32 to .72 for the PAS, and from .47 to .69 for the OAS.

Insert Table 1 about here

Insert Table 2 about here

Insert Table 3 about here

Criterion-related Validity

Criterion-related validity reveals how well scores on an instrument either predict future performance (i.e., predictive validity) or estimate current performance on another instrument that is hypothesized to measure a similar construct (i.e., concurrent validity). This evidence of validity is determined by relating performance on a test to performance on another criterion (Campbell & Stanley, 1990; Gay, 1999; Kerlinger, 1999). Evidence of concurrent validity

was established in the present study via significant correlations ($p < .001$) between scores on the Foreign Language Classroom Anxiety Scale (Horwitz et al., 1986) and scores on the Input Anxiety Scale, the Processing Anxiety Scale, and the Output Anxiety Scale. These correlations are presented in Table 4.

Insert Table 4 about here

The correlations between scores on the Foreign Language Classroom Anxiety Scale and scores on the Input Anxiety Scale, the Processing Anxiety Scale, and the Output Anxiety Scale in Table 4 are very similar in magnitude to those reported by MacIntyre and Gardner (1994a) (c.f., *Instruments and Procedure* section above). Indeed, transforming the correlations in both studies into Fisher's *z*-scores yielded no significant difference ($p < .05$) in magnitude between the correlations reported in Table 4 and the corresponding correlations in MacIntyre and Gardner's (1994a) study.

Invariance of Scales

Descriptive statistics were computed for each scale (range = 6 - 30). The mean for the IAS was 18.56 ($SD = 4.04$), for the PAS, 17.80 ($SD = 4.06$), and for the OAS, 19.36 ($SD = 4.13$). A series of dependent *t*-tests[8], using the Bonferroni adjustment (Huberty, 1994), revealed that the OAS generated statistically significantly higher mean scores than did the IAS ($t = 3.5$, $df = 256$, $p < .001$) and the PAS ($t = 7.8$, $df = 256$, $p < .001$). Also, the IAS generated statistically significantly higher mean scores than did the PAS ($t = 3.5$, $df = 256$, $p < .001$). These findings indicate that students reported significantly higher levels of output anxiety than input anxiety and processing anxiety, and significantly higher levels of input anxiety than processing anxiety.

A Kruskal-Wallis one-way analysis of variance[9] revealed no difference among students enrolled in the four language areas (i.e., Spanish, French, German, and Japanese) with respect to scores on the IAS ($\chi^2 = 1.63$; df = 3; $p > 0.05$), PAS ($\chi^2 = 1.38$; df = 3; $p > 0.05$), and OAS ($\chi^2 = 1.33$; df = 3; $p > 0.05$).

Additionally, a series of analysis of variance (ANOVA) tests[10] was conducted using gender and course level as independent variables. With regard to input anxiety, no significant differences were found among students enrolled in the introductory, intermediate, and advanced courses ($F_{2, 252} = 2.45$, $p > 0.05$), or between males and females ($F_{1, 252} = 2.72$, $p > 0.05$), nor was a course level X gender interaction found ($F_{2, 217} = 2.66$, $p > 0.05$). With respect to processing anxiety, no significant differences were found among students enrolled in the introductory, intermediate, and advanced courses ($F_{2, 252} = 0.77$, $p > 0.05$), or between males and females ($F_{1, 252} = 1.50$, $p > 0.05$), nor was a course level X gender interaction found ($F_{2, 217} = 0.86$, $p > 0.05$). Finally, with regard to output anxiety, no significant differences were found among students enrolled in the introductory, intermediate, and advanced courses ($F_{2, 252} = 0.30$, $p > 0.05$), or between males and females ($F_{1, 252} = 2.94$, $p > 0.05$), nor was a course level X gender interaction found ($F_{2, 217} = 0.10$, $p > 0.05$). Finally, a Kruskal-Wallis one-way analysis of variance revealed no difference in input anxiety ($\chi^2 = 1.37$; df = 4; $p > 0.05$), processing anxiety ($\chi^2 = 7.47$; df = 4; $p > 0.05$), and output anxiety ($\chi^2 = 7.85$; df = 4; $p > 0.05$) between students in different years of study.

A multiple regression analysis was used to determine which of the three scales was the best predictor of global foreign language anxiety, as measured by the FLCAS. Specifically, *a hierarchical* regression (Tabachnick & Fidell, 1996) was utilized whereby the order of entry of variables into the model reflected

MacIntyre and Gardner's (1994a) three-stage conceptualization. That is, the IAS was entered first into the model, followed by the PAS and the OAS.

The regression analysis revealed that all three scales contributed significantly ($F[3, 254] = 191.15$, $p < .0001$) to the prediction of global foreign language anxiety. These three scales together explained 69.4% of the variance in global anxiety (*adjusted* R^2 = 68.9%), suggesting a very large effect size (Cohen, 1988). The IAS (standardized beta coefficient = 0.18) made the biggest contribution, explaining 40.8% of the variance in global foreign language anxiety. With the inclusion of the IAS in the model, the PAS (standardized beta coefficient = 0.45) explained an additional 23.6% of the variance. The PAS (standardized beta coefficient = 0.32) accounted for a further 4.9% of the variance.

Multivariate Structure of the Three Scales

In order to assess simultaneously the structure of the three scales, a maximum likelihood confirmatory factor analysis was undertaken (Bollen, 1989). Three models representing alternative conceptualizations of the structure of these scales were tested. The first model hypothesized a single unidimensional factor underlying the IAS, the PAS, and the OAS. The extent to which this model is adequate justifies the combining of each scale's score to obtain a total score. In addition to the one-factor model, two full three-factor models were evaluated, comprising a full three-factor model in which the three factors were orthogonal (i.e., an orthogonal model) and a full three-factor model in which the factors were related (i.e., an oblique model). The latter model, namely the full three-factor oblique model, assumed that the three scales adequately measure and reflect MacIntyre and Gardner's (1994a) three-stage conceptualization of foreign language anxiety. That is, the full three-factor oblique model assumed that the

three scales represented three distinct but related constructs, and thus was the model of primary interest.

The following indices were used as measures of model fit: Chi-square (χ^2), the ratio of Chi-square to degrees of freedom (χ^2/df), and the Adjusted Goodness-of-Fit Index. Also, an independence model was tested to allow computation of the relative fit index (RFI), the incremental fit index (IFI), the Tucker-Lewis index (TLI), and the comparative fit index (CFI) (Bentler, 1990; Bentler & Bonett, 1980; Bollen, 1986, 1989; Schumaker & Lomax, 1996).

Results of the application of the alternative models are presented in Table 5. The independence model, composed of 18 independent factors (i.e., each item of each scale represented a factor), provided a poor fit to the data. The one-factor model, although providing substantial improvement over the independence model, also was inadequate as a representation of the simultaneous structure of the three scales. The full three-factor orthogonal model also provided substantial improvement over the independence model. However, this model was inferior to the one-factor model. Finally, the full three-factor oblique model was a considerable improvement over the full three-factor orthogonal model, the single-factor model, and the independence model. Nevertheless, the chi-square was still statistically significant, suggesting an inadequate fit (although it should be noted that sample sizes that exceed 200, as in the present study, tend to increase the likelihood that the chi-square test will indicate a significant probability level) (Schumaker & Lomax, 1996, p. 125). Furthermore, although the χ^2/df ratio of 2.63 is within the range of between 2 to 1 and 3 to 1 recommended by some researchers (e.g., Carmines & McIver as cited in Arbuckle, 1997) for declaring an acceptable fit, most researchers (e.g., Byrne, 1989) believe that relative chi-square ratios above 2.00 represent an inadequate fit. Thus, the

BEST COPY AVAILABLE

χ^2/df ratio in the present study was considered too high to justify declaring that the full three-factor oblique model fit the data. The goodness of fit index (GFI) and the adjusted goodness of fit index (AGFI), although much larger than that for the competing models, was smaller than the commonly used cut-off of .9 for deeming a model to be acceptable (Bentler & Bonett, 1980; Hu & Bentler, 1995). However, it could be argued that the GFI and the AGFI are relatively close to this cut-off point.

Insert Table 5 about here

The root mean square error of approximation (RMSEA; Browne & Cudeck, 1993), which is the square root of the mean squared difference between the original and the reproduced correlation matrix, is used to compare the fit of two different models to the same data. Browne and Cudeck (1993) assert that (1) a RMSEA of approximately .05 or less is indicative of a close fit of the model in relation to the degrees of freedom, (2) a RMSEA value between .05 and .08 indicates a reasonable error of approximation, and (3) models with RMSEA's greater than 0.1 always should be rejected. The value of 0.08 (90% confidence interval is .07 to .09) in Table 5 thus suggests that the full three-factor model can perhaps be improved.

The following indices were computed for comparison of the one-factor model, the full three-factor orthogonal model, and the full three-factor oblique model to the independence model: Bentler and Bonett's (1980) normed fit index (NFI), Bollen's (1986) relative fit index (RFI), Bollen's (1989) incremental fit index (IFI), the Tucker-Lewis index (TLI; Bentler & Bonett, 1980), and Bentler's (1990) comparative fit index (CFI). Using a cut-off of .90 (Bentler & Bonett, 1980),

it can be seen that the values pertaining to the full three-factor oblique model presented in Table 5 fall slightly short. These indices combined suggest that the full three-factor oblique model may not be an adequate explanation of the data.

The Pearson product-moment correlations between the factors pertaining to the full three-factor oblique model are presented in Table 6. All correlations between factors were statistically significant. Interestingly, the PAS and OAS factors were strongly related, raising an issue concerning their separation as constructs.

Insert Table 6 about here

Table 7 presents the unstandardized factor loadings, the standard errors pertaining to the unstandardized factor loadings, the large sample *t*-values for each unstandardized factor loading, and the standardized factor loadings. It can be seen from this table that, after the Bonferroni adjustment for Type I error is made, all factor loadings remained statistically significant. However, it is commonly recommended (e.g., Hatcher, 1994) that standardized factor loadings be interpreted alongside unstandardized factor loadings. Table 7 reveals that one item (i.e., Item 2 of the IAS) had a loading less than .3, three items had loadings between .3 and .4, two items had loadings between .4 and .5, four items had loadings between .5 and .6, four items had loadings between .6 and .7, two items had loadings between .7 and .8, and two items had loadings of .80 or greater. All the standardized factor loadings, except Item 2 of the IAS, exceeded .3. Whereas some researchers use Lambert and Durand's (1975) cut-off of .3 for deeming a factor loading noteworthy, others (e.g., Hatcher, 1994) contend that

a cut-off of .6 should be utilized. In any case, it is clear that some items (e.g., Items 4 and 6 of the IAS) loaded more strongly on their factors than did others (e.g., Item 2 of the IAS). Thus, the following three follow-up confirmatory factor analyses were undertaken: (1) a three-factor oblique model eliminating items with loadings less than .4; (2) a three-factor oblique model eliminating items with loadings less than .5; and (3) a three-factor oblique model eliminating items with loadings less than .6.

Insert Table 7 about here

The results of the application of these three additional models are presented in Table 8. The first model, namely, the three-factor oblique model containing items greater than or equal to .4, excluded the following four items: (1) Item 2 of the IAS (i.e., "It does not bother me if my French/Spanish/German/Japanese notes are disorganized before I study them"); (2) Item 3 of the IAS (i.e., "I enjoy just listening to someone speaking French/Spanish/German/Japanese"); (3) Item 3 of the PAS (i.e., "The only time that I feel comfortable during French/Spanish/German/Japanese tests is when I have had a lot of time to study"); and (4) Item 4 of the PAS (i.e., "I feel anxious if French/Spanish/German/Japanese class seems disorganized"). Thus, the three-factor oblique model containing items greater than or equal to .4 comprised 4 IAS items, 4 PAS items, and 6 OAS items. This model was an improvement over the full three-factor oblique model containing all items (see Table 5), as well as the other previous models (i.e., the full three-factor orthogonal model, the single-factor model, and the independence model). Although the chi-square was still statistically significant, the GFI and the AGFI were larger than those for

the competing models, though still slightly smaller than the cut-off of .9. Also, the NFI, RFI, IFI TLI, and CFI were all larger than those for the full three-factor oblique model. Indeed, these indices ranged from .80 to .88--close to an adequate fit.

Insert Table 8 about here

The three-factor oblique model containing items greater than or equal to .5 excluded the four items eliminated from the three-factor oblique model containing items greater than or equal to .4, as well as two additional items: (1) Item 1 of the IAS (i.e., "I am not bothered by someone speaking quickly in French/Spanish/German/Japanese"); and (2) Item 5 of the OAS (i.e., "I never get nervous when writing something for my French/Spanish/German/Japanese class"). Thus, the three-factor oblique model containing items greater than or equal to .5 comprised 3 IAS items, 4 PAS items, and 5 OAS items. This model was an even further improvement than its predecessor (Table 5). Again, the chi-square was statistically significant. However, all the fit indices approached .9, suggesting an acceptable fit.

Finally, the three-factor oblique model containing items greater than or equal to .6 excluded the six items eliminated from the three-factor oblique model containing items greater than or equal to .5, as well as four additional items: (1) Item 5 of the IAS (i.e., "I get upset when I read in French/Spanish/German/Japanese because I must read things again and again"); (2) Item 5 of the PAS (i.e., "I am self-confident in my ability to appreciate the meaning of French/Spanish/German/Japanese dialogue"); (3) Item 1 of the OAS (i.e., "I never feel tense when I have to speak in

French/Spanish/German/Japanese"); and (4) Item 4 of the OAS (i.e., "I never get nervous when writing something for my French/Spanish/German/Japanese class"). Thus, the three-factor oblique model containing items greater than or equal to .6 comprised 2 IAS items, 3 PAS items, and 3 OAS items. This model was found to provide the most adequate fit to the data. Most of the fit indices were greater than .9.

Discussion

Anxiety has been found to play a central role in the foreign language learning context (e.g., Onwuegbuzie et al., 1999b). Thus, the purpose of the present study was to examine the psychometric properties of the Input Anxiety Scale, the Processing Anxiety Scale, and the Output Anxiety Scale--measures of anxiety at three different stages of the foreign language learning process. Apart from MacIntyre and Gardner (1994a), no other study has examined the psychometric qualities of these instruments.

When analyzed separately, all three scales were found to possess adequate psychometric characteristics. Evidence of structural validity was established via exploratory factor analysis, which revealed one specific factor for each scale, explaining a large proportion of the variance in IAS, PAS, and OAS scores. All six items loaded on their respective scales. Additionally, evidence of criterion-related validity, specifically, concurrent validity, was provided via significant correlations between scores on the three instruments and scores on the FLCAS, a measure of global foreign language anxiety. With respect to reliability, Cronbach's Coefficient Alphas and the Point Multi-Serial Correlation Alpha coefficients indicated that the items in each scale were homogeneous. All three scales were found to be invariant with respect to gender, year of study, type of language course, and level of language course. Students reported higher levels

of output anxiety than the other forms of anxiety. Interestingly, input anxiety was found to be the most closely related to global foreign language anxiety, explaining slightly more than 40% of the total variance in the latter.

Although the three scales appear to have adequate psychometric properties, the confirmatory factor analysis did not provide sufficient evidence that these scales, in their present form, adequately measure and reflect MacIntyre and Gardner's (1994) three-stage conceptualization of foreign language anxiety. Nevertheless, several reasons might explain why the confirmatory factor analysis did not support the full three-factor oblique model. First and foremost, as noted by Skehan (1991), the acceptance or rejection of a confirmatory factor model is not only a function of the difference between the model and reality, it also is a function of the size of the sample. In particular, large samples tend to have a bias toward rejection of models (Skehan, 1991). According to Schumaker and Lomax (1996, p. 125), for sample sizes larger than 200, as in the current study, "the χ^2 test has a tendency to indicate a significant level" and, consequently, to lead to a rejection of the underlying model. Thus, the present sample size may explain, at least in part, why the full three-factor model was rejected.

Yet, it should be noted that, in addition to χ^2 values, various effect size indices were reported which strengthened the rationale for rejecting the full three-factor oblique model. Notwithstanding, several Monte Carlo studies (i.e., studies in which a series of specific empirical sampling distributions for each index are examined) have demonstrated that many of these indices also are affected by sample size. For example, Marsh, Balla, and McDonald (1988), who analyzed the distributions of 29 different indices (e.g., GFI, NFI, TLI), found several of these indices to be related to sample size. Notwithstanding, in most cases, all the fit indices obtained using ML techniques, the method used in the

present study, tend to perform much better with respect to accuracy of estimates and correctness of statistical results than those obtained using other techniques such as generalized least squares and the asymptotic distribution free method (Hu & Bentler, 1995). Regardless, it is clear that a replication of this study is needed using a range of sample sizes.

Apart from sample bias, violation of assumptions underlying estimation methods--specifically, violation of distributional assumptions and the effect of dependence of latent variates--can threaten the adequacy of fit indices. In particular, Hu and Bentler (1995) reported that, when latent variables are dependent, most fit indices over-reject models at a sample size of 250 or less. Interestingly, the present sample size of 258 students is very close to this cut-off point. Even more importantly, although foreign language anxiety has been conceptualized as occurring at three stages (MacIntyre & Gardner, 1994a), the fact that these stages are somewhat interdependent (MacIntyre & Gardner, 1994a) makes it likely that the latent variables are dependent. Indeed, the intercorrelations of the IAS, PAS, and OAS (Table 4) were large. This dependency among the latent variables might explain why the model was rejected. Given that chi-square tests have a tendency to reject models using sample sizes greater than 200, and that most fit indices lead to an over-rejection of models for samples smaller than 250 when latent variables are dependent, it is difficult, if not impossible, to recommend an ideal sample size for future replication studies.

It should be noted that the three measures of foreign language anxiety each each contain six items, which could be considered relatively few. It is possible that this small number of items reduced the fit indices, since the goodness of fit of a more parameterized model tends to be greater than that for simpler models because of the loss of degrees of freedom associated with the more complex

BEST COPY AVAILABLE

model (Mulaik, 1990). Thus, increasing the number of items in each scale may not only improve the psychometric properties of these scales, but also may improve the adequacy fit of the three-factor model.

Interestingly, however, the standardized factor loadings led to the identification of items which reduced the ability of the full three-factor oblique model to fit the data. In the absence of these problem items, the fit of the data improved substantially. Thus, it appears that these items should either be modified or discarded. The question is, how many items were problematic? When 4 items were eliminated, the model fit was marginal. When 6 items were discarded, the fit was adequate. Finally, when 10 items were removed, the fit was good. Future research should investigate further the optimal number of items to be modified/removed. One approach could be to begin by modifying the four items that had standardized factor loadings less than .4. These items involved two IAS items and two PAS items. Interestingly, two of these items pertained to the anxiety arising from feelings of disorganization. Indeed, these two items had the smallest factor loadings that emerged from the exploratory factor analyses (see Tables 1-3). Thus, it possible that feelings of disorganization lead to relatively ambivalent responses with respect to levels of anxiety. As such, perhaps, these two items should be discarded or replaced rather than modified.

In any case, once the first round of revisions are made, the three measures should then be re-administered, and the responses re-analyzed along the lines outlined in the current paper. This process should continue until the scales possess adequate psychometric properties both at the unidimensional and multidimensional levels.

Taken together, the findings of this study provide evidence that the IAS, PAS, and OAS, when used in a univariate manner, appear to generate reliable and

valid scores. Unfortunately, the multidimensional structure of these scales is in question. Nevertheless, the fact that an adequate fit was obtained when some items were eliminated suggests that careful refinement of these scales may result in the firm support of MacIntyre and Gardner's (1994a) theory that foreign language anxiety occurs at the input, processing, and output stages of the second language acquisition process. Indeed, the authors currently are using item response theory (i.e., Rasch one-parameter modeling) to investigate the hierarchical structure of the IAS, PAS, and OAS items. It is hoped that such research will lead to measures of anxiety at the three different stages of second language acquisition that could be used for diagnostic purposes, which, in turn, would help to increase our understanding about foreign language anxiety.

Notes

1. The authors contributed equally to this article.

2. The authors wish to acknowledge the Research Council of the University of X which provided funding for this project. In addition we wish to express our sincere appreciation to the faculty of the Department of Foreign Languages who assisted in data collection.

3. See also Campbell and Ortiz, 1991; Daly, 1991; MacIntyre and Gardner, 1994b; Phillips, 1992; Powell, 1991; Price, 1991; and Young, 1991.

4. See also Gardner, Smythe, and Lalonde 1984; Horwitz et al., 1986; and Phillips, 1992.

5. Tobias (1986) himself cautions that his model "arbitrarily separates the instructional process into the three classical information-processing components: input, processing, and output" (p. 36).

6. PMSCACs are different than item-total correlations. Whereas the PMSCACs represent alpha coefficients that are computed for the scale after the corresponding item has been removed, an item-total correlation represents the correlation between the response made to an item by each individual and his/her corresponding total scores for that scale to which the item belongs. The major difference between the two indices is that, whereas a PMSCAC helps to determine what happens to the overall internal consistency of a scale when an item is deleted, an item-total correlation indicates the extent to which a person's response to a particular item is predictive of her/his average response to all items. Although PMSCACs and item-total correlations yield different scores, they are often similar. Thus, typically it is redundant to report both indices.

7. Indeed, it is commonly argued that a principal components analysis should not be used to identify the number and nature of the factors that are responsible for covariation in the dataset because it makes no attempt to separate the common component from the unique component of each variable's variance. Thus, principal components analysis can provide a misleading representation of the factor structure underlying the data. For more information about the difference between factor analysis and principal components analysis, see Hatcher (1994).

8. Although some researchers undertake one-way repeated measures analyses of variance (ANOVAs) in order to determine whether there are statistically significant differences among multiple measures (i.e., an omnibus test), and then, if a significant difference is found, follow up with a series of α-protected (e.g., Scheffé tests) univariate analyses, this practice is now outdated. Moreover, many statisticians criticize this technique because analyses involving repeated measures test "linear combinations of the outcome variables (determined by the variable intercorrelations) and therefore do not

yield results that are in any way comparable with a collection of separate univariate tests" (Keselman et al., 1998 p. 361).

9. The Kruskal-Wallis test is the most powerful nonparametric test for examining three or more independent groups. It has 95 percent of the power of the *F* statistic (i.e., ANOVA) to detect existing differences between groups. This technique tests the null hypothesis that all samples are from the same population. In this study, the Kruskal-Wallis test was used to compare the language groups, instead of the parametric analysis of variance test (ANOVA), because the number of Japanese students (n = 5) was small, and thus a normal distribution could not be assumed for their anxiety scores. For a further discussion of use and interpretation of Kruskal-Wallis tests, the reader is referred to Hollander and Wolfe (1973).

10. A Multiple Analysis of Variance (MANOVA) followed by appropriate univariate analyses (i.e., a MANOVA-univariate data analysis strategy) was not conducted because "there is very limited empirical support for this strategy" (Keselman et al., 1998, p. 361). Indeed, Keselman et al. (1998) states that "If the univariate effects are those of interest, then it is suggested that the researcher go directly to the univariate analyses and bypass MANOVA....Focusing on results of multiple univariate analyses preceded by a MANOVA is no more logical than conducting an omnibus ANOVA but focusing on the results of group contrast analyses (Olejnik & Huberty, 1993)" (pp. 361-362). For a more extensive discussion of MANOVA versus multiple ANOVAs, see Huberty & Morris, 1989).

References

Aida, Y. (1994). Examination of Horwitz, Horwitz, and Cope's construct of foreign language anxiety: The case of students of Japanese. *The Modern Language Journal, 78,* 155-168.

Arbuckle, J.L. (1997). *AMOS Users' Guide Version 3.6.* Chicago, IL: SmallWaters Corporation.

Bailey, P., Onwuegbuzie, A.J., & Daley, C.E. (in press). Study habits and anxiety about learning foreign languages. *Perceptual and Motor Skills.*

Bentler, P.M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107,* 238-246.

Bentler, P.M., & Bonett, D.G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin, 88,* 588-606.

Bickel, P.J., & Doksum, K.A. (1977). *Mathematical statistics.* San Francisco: Holden-Day.

Bollen, K.A. (1986). Sample size and Bentler and Bonett's nonnormed fit index, *Psychometrika, 51,* 375-377.

Bollen, K.A. (1989). *Structure equations with latent variables.* New York: Wiley.

Browne, M.W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K.A. Bollen, & J.S. Long (Eds.), *Testing structural equation models* (pp. 136-162). Newbury Park, California: Sage.

Byrne, B.M. (1989). A primer of LISREL: *Basic applications and programming for confirmatory factor analytic models.* New York: Springer-Verlag.

Campbell, C.M., & Ortiz, J. (1991). Helping students overcome foreign language anxiety: A foreign language anxiety workshop. In E.K. Horwitz & D. J.

Young (Eds.), *Language anxiety: From theory and research to classroom implications* (pp. 153-168). Englewood Cliffs, NJ: Prentice-Hall.

Campbell, D.T., & Stanley, J.C. (1990). Experimental and quasi-experimental designs for research. Boston, MA: Houghton Mifflin Company.

Chastain, K. (1975). Affective and ability factors in second language acquisition. *Language Learning, 25,* 153-161.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). New York: John Wiley.

Daly, J. (1991). Understanding communication apprehension: An introduction for language educators. In E.K. Horwitz and D. J. Young, (Eds.), *Language anxiety: From theory and research to classroom implications,* (pp. 3-13). Englewood Cliffs, NJ: Prentice-Hall.

Ehrman, M., & Oxford, R. (1995). Cognition plus: Correlates of language learning success. *Modern Language Journal,* 79, 67-89.

Ganschow , L. & Sparks, R. (1996). Anxiety about Foreign Language Learning among High School Women. *The Modern Language Journal, 80,* 199-212.

Gardner, R.C. (1985). *Social psychology and second language learning: The role of attitudes and motivation.* London: Edward Arnold.

Gardner, R.C., & MacIntyre, P.D. (1993). On the measurement of affective variables in second language learning. *Language Learning, 43,* 157-194.

Gardner, R.C., Smythe, P.C., & Lalonde, R.N. (1984). *The nature and replicability of factors in second language learning* (Research Bulletin No. 605). London, Ontario: The University of Western Ontario.

Gay, L.R. (1999). *Educational research: Competencies for analysis and application* (6th ed.). New York: Merrill.

Hatcher, L. (1994). A step-by-step approach to using the SAS system for

factor analysis and structural equation modeling. Cary, NC: SAS Institute Inc..

Hollander, M., & Wolfe, D.A. (1973). Nonparametric statistical methods. New York: John Wiley & Sons, Inc.

Horwitz, E.K. (1986). Preliminary evidence for the reliability and validity of a foreign language anxiety scale. *TESOL Quarterly, 20,* 559-564.

Horwitz, E.K., Horwitz, M.B., & Cope J. (1986). Foreign language classroom anxiety. *Modern Language Journal, 70,* 125-132.

Hu, L.-T. & Bentler, P.M. (1993). *Fit indexes in covariance structural equation modeling.* Unpublished manuscript, University of California, Santa Cruz.

Hu, L.-T. & Bentler, P.M. (1995). Evaluating model fit. In R.H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (pp. 76-99). Thousand Oaks, CA: Sage Publications, Inc.

Huberty, C.J. (1994). *Applied discriminant analysis.* New York: Wiley.

Huberty, C.J., & Morris, J.D. (1989). Multivariate analysis versus multiple univariate analyses. *Psychological Bulletin, 105,* 302-308.

Kerlinger, F.N. (1999). Foundations of behavioral research (4th ed.). Fort Worth, Texas: Harcourt Brace College Publishers.

Keselman, H.J., Huberty, C.J., Lix, L.M., Olejnik, S., Cribbie, R.A., Donahue, B., Kowalchuk, R.K., Lowman, L.L., Petoskey, M.D., Keselman, J.C., & Levin, J.R. (1998). Statistical practices of educational researchers: An analysis of their ANOVA, MANOVA, and ANCOVA analyses. *Review of Educational Research, 68,* 350-386.

Lambert, Z.V. & Durand, R.M. (1975). Some precautions in using canonical analysis. *Journal of Market Research, XII,* 468-475.

Lawley, D.N., & Maxwell, A.E. (1971). *Factor analysis as a statistical method.* New York: Macmillan Publishing Co., Inc.

MacIntyre, P.D. (1999). Language anxiety: A review of the research for language teachers. In D.J. Young (Ed.), *Affect in foreign language and second language learning: A practical guide to creating a low-anxiety classroom atmosphere*. San Francisco: McGraw-Hill.

MacIntyre, P.D., & Gardner, R.C. (1988). *The measurement of anxiety and applications to second language learning: An annotated bibliography* (Department of Psychology Research Bulletin No. 672). London, Ontario: The University of Western Ontario. (ERIC Document Reproduction Service No. FL 017 649).

MacIntyre, P.D., & Gardner, R.C. (1991a). Language anxiety: Its relation to other anxieties and to processing in native and second languages. *Language Learning, 41*, 85-117.

MacIntyre, P.D., & Gardner, R.C. (1991b). Investigating language class anxiety using the focused essay technique. *The Modern Language Journal, 75*, 296-304.

MacIntyre, P.D., & Gardner, R.C. (1991c). Methods and results in the study of foreign language anxiety: A review of the literature. *Language Learning, 41*, 85-117.

MacIntyre, P.D., & Gardner, R.C. (1994a). The subtle effects of language anxiety on cognitive processing in the second language. *Language Learning, 44*, 283-305.

MacIntyre, P.D., & Gardner, R.C. (1994b). The effects of induced anxiety on three stages of cognitive processing in computerized vocabulary learning. *Studies in Second Language Acquisition*, 16: 1-17.

MacIntyre, P.D., Noels, K.A., & Clément, R. (1997). Biases in self-ratings of second language proficiency: The role of language anxiety. *Language Learning, 47*, 265-287.

Madsen, H.S., Brown, B.B., & Jones, R.L. (1991). Evaluating student attitudes toward second language tests. In E.K. Horwitz & D. J. Young (Eds.), *Language anxiety: From theory and research to classroom implications* (pp. 65-86). Englewood Cliffs, NJ: Prentice-Hall.

Marsh, H.W., Balla, J.R., & McDonald, R.P. (1988). Goodness-of-fit indexes in confirmatory factor analysis: The effect of sample size. *Psychological Bulletin, 103,* 391-410.

Mulaik, S.A. (1990, June). *An analysis of the conditions under which the estimation of parameters inflates goodness-of-fit indices as measures of model validity.* Paper presented at the annual meeting of the Psychometric Society, Princeton, NJ.

Nunnally, J.C. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.

Olejnik, S., & Huberty, C.J. (1993, April). *Preliminary statistical tests.* Paper presented at the annual meeting of the American Educational Research Association, Atlanta, GA.

Onwuegbuzie, A.J., Bailey, P., & Daley, C.E. (1999a). *Cognitive, affective, personality, and demographic predictors of foreign language achievement.* Unpublished manuscript, Valdosta State University, Valdosta, GA.

Onwuegbuzie, A.J., Bailey, P., & Daley, C.E. (1999a). *The role of foreign language anxiety and students' expectations in foreign language learning.* Unpublished manuscript, Valdosta State University, Valdosta, GA.

Onwuegbuzie, A.J., Bailey, P., & Daley, C.E. (in press). Factors associated with foreign language anxiety. *Applied Psycholinguistics.*

Onwuegbuzie, A.J., & Daley, C.E. (1996). The relative contributions of examination-taking coping strategies and study coping strategies on test anxiety:

A concurrent analysis. *Cognitive Therapy and Research, 20,* 287-303.

Phillips, E.M. (1992). The effects of language anxiety on student oral test performance and attitudes. *The Modern Language Journal, 76,* 14-26.

Powell, J.A. (1991). Foreign language classroom anxiety: Institutional response. In E.K. Horwitz & D.J. Young (Eds.), *Language anxiety: From theory and research to classroom implications* (pp. 169-176). Englewood Cliffs, NJ: Prentice Hall.

Price, M.L. (1991). The subjective experience of foreign language anxiety: Interviews with highly anxious students. In E.K. Horwitz & D.J. Young (Eds.), Language anxiety: From theory and research to classroom implications (pp. 101-108). Englewood Cliffs, NJ: Prentice-Hall.

Schumaker, R.E., & Lomax, R.G. (1996). *A beginner's guide to structural equation modeling*. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Skehan, P. (1991). Individual differences in second language learning. *Studies in Second Language Acquisition, 13,* 275-298.

Steinberg, F.S., & Horwitz, E.K. (1986). The effect of induced anxiety on the denotative and interpretive content of second language speech. *TESOL Quarterly, 20,* 131-136.

Tabachnick, B.G., & Fidell, L.S. (1996). *Using multivariate statistics*. (3rd ed.). New York: HarperCollins College Publishers.

Tobias, S. (1977). A model for research on the effect of anxiety on instruction. In Sieber, J., O'Neil. H.F., Jr., & Tobias, S. *Anxiety, learning and instruction*. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Tobias, S. (1986). Anxiety and cognitive processing of instruction. In Schwarzer, R. (Ed.), *Self-related cognitions in anxiety and motivation*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Young, D.J. (1991). Creating a low anxiety classroom environment: What does language anxiety research suggest? *The Modern Language Journal*, 75, 426-439.

Table 1:

Factor Loadings and Percents of Variance for One-Factor Common Factor Analysis on IAS Items (N = 258)

Item		Factor Loading	Point Multi-Serial Coefficient Alpha
1	I am not bothered by someone speaking quickly in French/ Spanish/German/Japanese.	.49*	.69
2.	It does not bother me if my French/Spanish/German/Japanese notes are disorganized before I study them.	.30*	.74
3.	I enjoy just listening to someone speaking French/Spanish/ German/Japanese.	.42*	.71
4.	I get flustered unless French/Spanish/German/Japanese is spoken very slowly and deliberately.	.77*	.62
5.	I get upset when I read in French/Spanish/German/ Japanese because I must read things again and again.	.57*	.68
6.	I get upset when French/ Spanish/German/Japanese is spoken too quickly.	.78*	.62

% of total variance accounted for by the solution = 43.3

* loadings with large effect sizes, using a cut-off loading of 0.3 recommended by Lambert and Durand (1975)

Table 2:

Factor Loadings and Percents of Variance for One-Factor Common Factor Analysis on PAS Items (N = 258)

Item		Factor Loading	Point Multi-Serial Coefficient Alpha
1	Learning new French/Spanish/German/ Japanese vocabulary does not worry me, I can acquire it in no time.	.68*	.67
2.	I am anxious with French/Spanish/ German/Japanese because, no matter how hard I try, I have trouble understanding it.	.66*	.67
3.	The only time that I feel comfortable during French/Spanish/ German/Japanese tests is when I have had a lot of time to study.	.50*	.72
4.	I feel anxious if French/Spanish/ German/Japanese class seems disorganized.	.32*	.75
5.	I am self-confident in my ability to appreciate the meaning of French/Spanish/ German/Japanese dialogue.	.50*	.69
6.	I do not worry when I hear new or unfamiliar words, I am confident that I can understand them.	.72*	.65

% of total variance accounted for by the solution = 44.0

* loadings with large effect sizes, using a cut-off loading of 0.3 recommended by Lambert and Durand (1975)

Table 3:

Factor Loadings and Percents of Variance for One-Factor Common Factor Analysis on OAS Items (N = 258)

Item		Factor Loading	Point Multi-Serial Coefficient Alpha
1	I never feel tense when I have to speak in French/Spanish/ German/Japanese.	.56*	.72
2.	I feel confident that I can easily use the French/Spanish/ German/Japanese vocabulary that I know in a conversation.	.56*	.72
3.	I may know the proper French/Spanish German/Japanese expression but when I am nervous it just won't come out.	.69*	.69
4.	I get upset when I know how to communicate in French/Spanish/ German/Japanese but I just cannot verbalize it.	.57*	.71
5.	I never get nervous when writing something for my French/Spanish/German/ Japanese class.	.47*	.74
6.	When I become anxious during a French/Spanish/German/Japanese test, I cannot remember anything I studied.	.63*	.71

% of total variance accounted for by the solution = 44.7

* loadings with large effect sizes, using a cut-off loading of 0.3 recommended by Lambert and Durand (1975)

Table 4:

Pearson Product-Moment Correlations Among IAS, PAS, OAS, and FLCAS (N = 258)

	IAS	PAS	OAS
1. IAS			
2. PAS	.61*		
3. OAS	.58*	.68*	
4. FLCAS	.64*	.77*	.73*

* $p < .001$

Table 5:

Goodness of Fit for Alternative Models

				Index							
Model	df	χ^2	χ^2/df	GFI	AFGI	RMSEA	NFI	RFI	IFI	TLI	CFI
Independence	153	1564.31*	10.22	.39	.31	.18	-	-	-	-	-
One-Factor Model	135	421.91*	3.13	.83	.79	.09	.73	.69	.80	.77	.80
Three-Factor Orthogonal Model	135	654.71*	4.85	.79	.73	.12	.58	.53	.64	.58	.63
Three-Factor Oblique Model	132	346.48*	2.63	.86	.82	.08	.78	.74	.85	.82	.85

* $p > .05$

df = degrees of freedom
GFI = goodness of fit index
AGFI = adjusted goodness of fit index
RMSEA = root mean square error of approximation
NFI = normed fit index
RFI = relative fit index
IFI = incremental fit index
TLI = Tucker-Lewis index
CFI = comparative fit index

Table 6:

Pearson Product-Moment Correlations Among Factors Pertaining to the Full Three-Factor Oblique Model

	1	2
1. IAS		
2. PAS	.78*	
3. OAS	.74*	.93*

* $p < .001$

Table 7:

Three-Factor Oblique Model: Unstandardized Factor Loadings, Standard Errors, Large Sample t-Values, and Standardized Factor Loadings for IAS, PAS, and OAS Items

Scale	Item No.	Unstandardized Factor Loading	Standard Error	*t*-Value	Standardized Factor Loading
IAS:					
	1	.56	.073	7.74'	.49''
	2	.29	.067	4.35'	.29
	3	.35	.063	5.47'	.36''
	4	.85	.059	14.29'	.80''
	5	.55	.057	9.69'	.59''
	6	.89	.061	14.63'	.81''
PAS:					
	1	.75	.065	11.59'	.68''
	2	.72	.059	12.25'	.71''
	3	.42	.068	6.14'	.39''
	4	.34	.069	4.94'	.32''
	5	.58	.062	9.47'	.58''
	6	.68	.056	12.26'	.71''
OAS:					
	1	.55	.058	9.43'	.58''
	2	.69	.061	11.32'	.67''
	3	.62	.061	10.07'	.61''
	4	.48	.061	7.91'	.50''
	5	.50	.070	7.12'	.46''
	6	.68	.068	10.01'	.61''

' significant after the Bonferroni adjustment for Type I error.

'' loadings greater than Lambert and Durand's (1975) cut-off value of 0.3.

Note: The item numbers correspond to those in Tables 1-3.

Table 8:

Goodness of Fit for Additional Models

		Index								
Model	df	χ^2	χ^2/df	GFI	AFGI	RMSEA	NFI	RFI	IFI	TLI	CFI
Three-Factor Oblique Model with Items with Loadings ≥ .4	74	220.29*	2.98	.88	.83	.09	.83	.80	.88	.86	.88
Three-Factor Oblique Model with Items with Loadings ≥ .5	51	177.73*	3.49	.89	.83	.10	.85	.81	.89	.85	.89
Three-Factor Oblique Model with Items with Loadings ≥ .6	17	64.86*	3.82	.94	.87	.11	.92	.86	.94	.89	.94

* $p > .05$

df = degrees of freedom
GFI = goodness of fit index
AGFI = adjusted goodness of fit index
RMSEA = root mean square error of approximation
NFI = normed fit index
RFI = relative fit index
IFI = incremental fit index
TLI = Tucker-Lewis index
CFI = comparative fit index